CONVERGENCE OF AN ITERATIVE METHOD FOR TOTAL VARIATION DENOISING

DAVID C. DOBSON* AND CURTIS R. VOGEL†

Abstract. In total variation denoising, one attempts to remove noise from a signal or image by solving a nonlinear minimization problem involving a total variation criterion. Several approaches based on this idea have recently been shown to be very effective, particularly for denoising functions with discontinuities. This paper analyzes the convergence of an iterative method for solving such problems. The iterative method involves a "lagged diffusivity" approach in which a sequence of linear diffusion problems are solved. Global convergence in a finite dimensional setting is established, and local convergence properties, including rates and their dependence on various parameters, are examined.

Key Words. denoising, total variation, convergence analysis.

AMS(MOS) subject classification. 49M05, 65K10

1. Introduction.

Consider the problem of reconstructing an unknown function u, called the image, from data z satisfying

$$z = u + \epsilon. \tag{1.1}$$

Here ϵ represents error, or noise, in the data. It is assumed that u is defined on a region $\Omega \subset R^d$, $d = 1, 2$, or 3, which is bounded and convex with Lipschitz continuous boundary $\partial\Omega$. The Euclidean norm on R^d will be denoted by $|\cdot|$.

A variety of linear filtering techniques may be applied. These tend to perform poorly when u has discontinuities or steep gradients. See Figure 1 of [17]. When u is blocky, i.e., nearly piecewise continuous, techniques based on total variation

$$TV(u) \stackrel{\text{def}}{=} \int_\Omega |\nabla u| \tag{1.2}$$

have been shown to be very effective. See [15, 17] for numerical evidence and [7] for a theoretical analysis.

Several approaches may be taken to incorporate total variation. Rudin, Osher, and Fatemi [15] considered the constrained minimization problem

$$\min_u TV(u) \quad \text{subject to} \quad \|u - z\|^2_{L^2(\Omega)} = \sigma^2, \tag{1.3}$$

where σ quantifies the error level and is assumed to be known. In the context of parameter identification, Gutman [12] also used constrained minimization, but with the constraint applied instead to $TV(u)$. A similar inverse problem with a linearized constraint was formulated by Dobson and Santosa in [8]. Total variation has also been

* Department of Mathematics, Texas A&M University, College Station, TX 77843-3368. dobson@math.tamu.edu. Work partially supported by an NSF Mathematical Sciences Postdoctoral Fellowship.

† Department of Mathematical Sciences, Montana State University, Bozeman, MT 59717-0240. Work partially supported by the NSF under grant DMS-9303222. vogel@math.montana.edu

successfully applied to deblurring problems in image processing. In this case, u in (1.1) is replaced by Ku, where K denotes a linear smoothing operator. For example, see [16, 18, 14].

Rather than explicitly enforcing constraints, one may also consider the penalized least squares problem

$$\min_u \frac{1}{2}\|u - z\|^2_{L^2(\Omega)} + \alpha\, TV(u), \tag{1.4}$$

where α is a positive parameter controlling the trade-off between goodness of fit to the data and variability in u. In the inverse and ill-posed problems community, this approach is known as Tikhonov Regularization.

Given an approach like (1.3) or (1.4), a variety of numerical solution techniques may be applied. To deal with the non-differentiability of the TV functional, interior point methods from linear programming (see Li and Santosa [14]; note that $TV(u)$ has been replaced by the ℓ_1 norm of the gradient of u) and augmented Lagrangian methods (see Ito and Kunisch [13]) have been proposed.

One may also replace the TV functional by a smooth approximation like

$$J(u) = \int_\Omega \sqrt{|\nabla u|^2 + \beta^2} \tag{1.5}$$

and then apply conventional optimization techniques–e.g., gradient descent [15] or Newton's method [17]. Both methods may require "globalization" (e.g., line search or some other step size control to guarantee convergence). In addition, gradient descent tends to converge very slowly. Global convergence can also be obtained by combining Newton's method with continuation in the parameter β. See the paper by T. Chan, Zhou, and R. Chan [2] for details.

In [17], Vogel and Oman introduced a fixed point iteration to minimize the penalized least squares functional

$$\frac{1}{2}\|u - z\|^2_{L^2(\Omega)} + \alpha J(u). \tag{1.6}$$

The corresponding Euler-Lagrange equations

$$u \; - \alpha \nabla \cdot \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \beta^2}} \right) = z, \quad \text{in } \Omega, \tag{1.7}$$

$$\frac{\partial u}{\partial n} = 0, \quad \text{on } \partial\Omega, \tag{1.8}$$

can be expressed in operator notation

$$(I + \alpha L(u))u = z, \tag{1.9}$$

where

$$L(u)v \stackrel{\text{def}}{=} -\nabla \cdot \left(\frac{1}{\sqrt{|\nabla u|^2 + \beta^2}}\, \nabla v \right). \tag{1.10}$$

The fixed point iteration is then

$$u^{\nu+1} = (I + \alpha L(u^\nu))^{-1} z \stackrel{\text{def}}{=} F(u^\nu), \quad \nu = 0, 1, \ldots . \tag{1.11}$$

Numerical experiments like those presented in [17] suggest that the convergence of (discretized versions of) this fixed point iteration is global and quite rapid. In addition, each fixed point iteration requires the inversion of discretizations of the positive, symmetric linear differential operator $I + \alpha L(u^\nu)$. This can be accomplished very efficiently using standard tools from numerical linear algebra, e.g., the conjugate gradient method with a multigrid preconditioner. See [17] for details.

The fixed point iteration (1.11) may be viewed as a special case of the "half quadratic regularization" scheme originated by D. Geman (see [9, 10]) and of the ARTUR algorithm of Charbonnier et al (see [4]). It is closely related to a "relaxation" algorithm proposed by Chambolle and Lions in [3].

The purpose of this paper is the analysis of discrete approximations of the fixed point iteration (1.11). In the next section we state some preliminary results concerning the well-posedness of the minimization problem and the consistency of discrete approximations. In Section 3 we introduce the fixed-point iteration in a finite dimensional setting and examine some of its properties. In Section 4, we establish the global convergence (i.e., convergence from *any* starting point) of the iteration—again in a finite dimensional setting. The analysis in this section mirrors that in [4] and is closely related to that in [3]. Finally, in Section 5, we examine some of the local convergence properties of iteration (1.11). In particular, we analyze the effects of various parameters (e.g., the "regularization" parameter α in (1.6), the "smoothness" parameter β in (1.5), and the discretization level) on rates of convergence.

2. Mathematical Preliminaries. Let $\beta \geq 0$ be fixed. Applying ideas in [1], let

$$\mathcal{V} = \{\vec{v} \in C_0^\infty(\Omega; \mathbb{R}^{\mathrm{d}}) : |\vec{\mathrm{v}}(\mathrm{x})| \leq 1\}, \tag{2.1}$$

and define the functionals $Q : L^2(\Omega) \times \mathcal{V} \to \mathbb{R}$

$$Q(u, \vec{v}) \stackrel{\text{def}}{=} \int_\Omega \left[-u \, \nabla \cdot \vec{v} + \beta \sqrt{1 - |\vec{v}|^2} \right] dx, \tag{2.2}$$

and $J : L^2(\Omega) \to \mathbb{R}$

$$J(u) \stackrel{\text{def}}{=} \sup_{\vec{v} \in \mathcal{V}} Q(u, \vec{v}). \tag{2.3}$$

If u is sufficiently smooth (e.g., in $C^1(\Omega)$), the functionals in (2.3) and (1.5) coincide. Moreover, $J(u)$ is convex and weakly lower semicontinuous. See [1] for details.

Let

$$f(u) \stackrel{\text{def}}{=} \frac{1}{2} ||u - z||^2 + \alpha J(u) \tag{2.4}$$

where α is a fixed positive parameter.

THEOREM 2.1. *There is a unique $u^* \in L^2(\Omega)$ for which*

$$f(u^*) = \inf_{u \in L^2(\Omega)} f(u). \tag{2.5}$$

Proof. Note that $f(u)$ is weakly lower semicontinuous and coercive. This combined with the weak compactness of closed balls in $L^2(\Omega)$ yields existence. Uniqueness follows from the strict convexity of the L^2 norm. $\square$

Let $\{\mathcal{U}^N\}$ be a sequence of subspaces of $L^2(\Omega)$ with associated (L^2) projection operators P_N.

THEOREM 2.2. *Let f_N denote the restriction of f to $\mathcal{U}^N$, and let $u_N \in \mathcal{U}^N$ denote the minimizer of f_N. Suppose that*

$$\liminf_{N \to \infty} f(P_N u^*) \leq f(u^*). \tag{2.6}$$

Then $u_N \to u^*$ as $N \to \infty$.

Proof. Note that

$$f(u_N) = f_h(u_N) \leq f_N(P_N u^*) = f(P_N u^*).$$

Now suppose that u_N does not converge to u^*. The u_N's are bounded in $L^2(\Omega)$, so there exists a weakly convergent subsequence $u_{N_j} \rightharpoonup \bar{u} \neq u^*$. From the weak lower semicontinuity of f and the above inequality,

$$f(\bar{u}) \leq \liminf f(u_{N_j}) \leq \liminf f(P_{N_j} u^*).$$

From the supposition (2.6), one obtains $f(\bar{u}) \leq f(u^*)$, which contradicts the uniqueness of the minimizer for f. $\square$

Condition (2.6) generally holds for "standard" finite element subspaces $\mathcal{U}^{\mathcal{N}}$, provided u^* is sufficiently smooth. This can be seen as follows. We have

$$\begin{aligned} f(P_N u^*) &= \frac{1}{2}\|P_N u^* - z\|_2^2 + \alpha J(P_N u^*) \\ &\leq f(u^*) + \alpha J(P_N u^* - u^*) + \mathcal{O}(\|P_N u^* - u^*\|_2) \\ &\leq f(u^*) + \mathcal{O}(\|P_N u^* - u^*\|_{W^{1,2}(\Omega)}). \end{aligned}$$

Then for example if $u^* \in H^2(\Omega) \cup W^{1,\infty}(\Omega)$ and $\mathcal{U}^N$ consists of continuous piecewise linear functions on a regular triangulation, one can show (by a slight modification of results in [5]) that

$$\|P_N u^* - u^*\|_{W^{1,2}(\Omega)} \leq Ch$$

where h represents maximum grid spacing. Approximation properties like this are well known for a wide variety of finite element spaces.

We note that a complete analysis of finite element approximations for a minimal surface problem closely related to the minimization of (2.4) is given in Ciarlet [5]. Our goal in this section was merely to show that this approximation is consistent. Careful study of discretization schemes for (2.4), particularly in the absence of stringent a priori smoothness assumptions on u^*, is an interesting topic for further research.

3. Fixed point iteration. In this section we introduce the fixed-point iteration (1.11) in a finite dimensional setting and examine some of its properties. We begin with a computation of Gateaux derivatives. Assume u, v are smooth, and define $\tilde{f}(\tau; u, v) = f(u + \tau v)$. Let $q(t) = \alpha\sqrt{t^2 + \beta^2}$. Then

$$q'(t) = \frac{\alpha}{\beta}\frac{t}{(1 + (t/\beta)^2)^{1/2}}$$

and

$$q''(t) = \frac{\alpha}{\beta}\frac{1}{(1 + (t/\beta)^2)^{3/2}}.$$

Hence, the first and second Gateaux derivatives of f are

$$\begin{aligned} \langle f'(u), v\rangle &\stackrel{\text{def}}{=} \frac{d\tilde{f}}{d\tau}(\tau; u, v)|_{\tau=0} \\ &= \int_\Omega (u(x) - z(x))\, v(x)\, dx + \frac{\alpha}{\beta}\int_\Omega \frac{\nabla u \cdot \nabla v}{\left(|\frac{\nabla u}{\beta}|^2 + 1\right)^{1/2}} \\ &\stackrel{\text{def}}{=} \langle ((1 + L(u))u - z)\,, v\rangle, \end{aligned} \tag{3.1}$$

and

$$\begin{aligned} \langle f''(u)v, v\rangle \stackrel{\text{def}}{=} \frac{d^2\tilde{f}}{d\tau^2}(\tau; u, v)|_{\tau=0} &= \int_\Omega v(x)^2\, dx + \frac{\alpha}{\beta}\int_\Omega \frac{|\nabla v|^2}{\left(|\frac{\nabla u}{\beta}|^2 + 1\right)^{3/2}} \\ &\stackrel{\text{def}}{=} \langle (1 + M(u))\, v, v\rangle. \end{aligned} \tag{3.2}$$

Thus, the gradient and Hessian of f are given, respectively, by

$$g(u) \stackrel{\text{def}}{=} (1 + L(u))\, u - z \tag{3.3}$$

and

$$H(u) \stackrel{\text{def}}{=} 1 + M(u) = 1 + L(u) + L'(u)u, \tag{3.4}$$

where

$$\langle L(u)v, v\rangle \stackrel{\text{def}}{=} \frac{\alpha}{\beta}\int_\Omega \frac{1}{\left(|\frac{\nabla u}{\beta}|^2 + 1\right)^{1/2}}|\nabla v|^2, \tag{3.5}$$

$$\langle (L'(u)u)\, v, v\rangle \stackrel{\text{def}}{=} -\frac{\alpha}{\beta^3}\int_\Omega \frac{|\nabla u|^2}{\left(|\frac{\nabla u}{\beta}|^2 + 1\right)^{3/2}}|\nabla v|^2. \tag{3.6}$$

At this point, to be concrete we introduce a particular (finite element) discretization of the underlying problem. We note however that the results following hold for any discretization which inherits certain monotonicity properties of the continuous operators. See for example the finite difference discretization in [15]. Consider the approximation

$$u \approx U \stackrel{\text{def}}{=} \sum_{j=1}^{N} U_j \phi_j, \tag{3.7}$$

where the U_j are scalars, and assume that

(**A.1**) the basis functions ϕ_j's are linearly independent;

(**A.2**) each $\phi_j \in W^{1,\infty}(\Omega)$.

Consequently,

$$U \in W^{1,\infty}(\Omega). \tag{3.8}$$

In an abuse of notation, U will represent both the coefficient vector $\{U_j\}_{j=1}^N$ and the corresponding linear combination (3.7). We then consider the discrete cost functional $\mathbb{R}^{\mathrm{N}} \to \mathbb{R}$ defined by

$$f(U) = \frac{1}{2}||U - z||^2_{L^2(\Omega)} + \alpha \int_\Omega \sqrt{|\nabla U|^2 + \beta^2}. \tag{3.9}$$

As a consequence of Theorem 2.1 and independence of the basis functions, the functional f has a unique minimizer U^*. The derivation above then naturally yields vectors $g(U) \in \mathbb{R}^N$ and matrices $L(U)$, $L'(U)U$ and $H(U)$ in $\mathbb{R}^{N\times N}$. It is readily observed from (3.5) that $L(U)$ is symmetric positive semidefinite and from (3.6) that $L'(U)U$ is symmetric negative semidefinite.

The fixed point iteration can be written

$$\begin{aligned} U^{\nu+1} &= (I + L(U^\nu))^{-1} z \\ &= U^\nu - (I + L(U^\nu))^{-1} g(U^\nu), \quad \nu = 0, 1, 2, \ldots. \end{aligned} \tag{3.10}$$

Thus

$$g(U^\nu) = -(I + L(U^\nu)) d^\nu, \tag{3.11}$$

where

$$d^\nu \stackrel{\text{def}}{=} U^{\nu+1} - U^\nu. \tag{3.12}$$

Note the similarity between the iteration (3.10) and Newton's method, where $(I + L(U^\nu))$ is replaced by $H(U^\nu)$.

The following properties are immediate consequences of (3.10)-(3.12) and the fact that $L(U^\nu)$ is positive semidefinite.

LEMMA 3.1. *For $\nu = 0, 1, 2, \ldots$,*

$$\langle g(U^\nu), d^\nu \rangle \leq -||d^\nu||^2, \tag{3.13}$$

$$||U^{\nu+1}|| \leq ||z||. \tag{3.14}$$

Property (3.13) implies that d^ν is a descent direction for f at U^ν, while (3.14) shows that the iterates are uniformly bounded in the L^2 norm. These are of course not sufficient to guarantee that the iteration converges. A proof of convergence in the finite dimensional setting is the topic of the next section. This proof will require a uniform H^1 bound on the iterates.

LEMMA 3.2. *There exists a constant $B > 0$ such that for $\nu = 1, 2, \ldots$,*

$$||U^\nu||_{H^1(\Omega)} \leq B. \tag{3.15}$$

Proof. Let $\mathbf{U} \in \mathbb{R}^{\mathrm{N}}$ denote the vector of coefficients in the representation (3.7). Then

$$||U||^2 = \int_\Omega U^2 = \mathbf{U}^T M_0 \mathbf{U}$$

and

$$||U||^2_{H^1(\Omega)} = \int_\Omega [|\nabla U|^2 + U^2] = \mathbf{U}^T M_1 \mathbf{U} + \mathbf{U}^T M_0 \mathbf{U},$$

where

$$[M_0]_{ij} = \int_\Omega \phi_i \, \phi_j, \quad [M_1]_{ij} = \int_\Omega \nabla \phi_i \cdot \nabla \phi_j.$$

From this and (3.14), equation (3.15) follows with

$$B = \sqrt{\rho(M_0^{-1}) \left(\rho(M_0) + \rho(M_1)\right)} \; ||z||. \tag{3.16}$$

Here $\rho(M)$ denotes the spectral radius of a matrix M. Note that because of assumption (**A.1**), M_0 is nonsingular. ◻

4. Global convergence. Adopting notation in Charbonnier et al [4], define

$$\begin{aligned}
\phi(t) &= \sqrt{t^2 + \beta^2} && (4.1)\\
\psi(s) &= \beta^2 s + \frac{1}{s} && (4.2)\\
\phi^*(t, s) &= \frac{1}{2}\left(s t^2 + \psi(s)\right) = \frac{1}{2}\left(s(t^2 + \beta^2) + \frac{1}{s}\right). && (4.3)
\end{aligned}$$

Fixing t, $0 \leq t < \infty$, setting $\frac{\partial \phi^*}{\partial s} = 0$, and noting that

$$\frac{\partial^2 \phi^*}{\partial s^2} = \frac{1}{s^3}, \tag{4.4}$$

which is positive for $0 < s < \infty$, one obtains

$$\phi(t) = \min_{s>0} \phi^*(t, s) = \phi^*(t, s^*) \tag{4.5}$$

where

$$s^* = \frac{1}{\sqrt{t^2 + \beta^2}}. \tag{4.6}$$

We shall again take a discretization (3.7) and assume that (**A.1**) and (**A.2**) hold, so that (3.8) is satisfied. The discrete cost functional (3.9) can then be expressed as

$$\begin{aligned}
f(U) &= \frac{1}{2}||U - z||^2 + \alpha \int_\Omega \phi(|\nabla U|) && (4.7)\\
&= \inf_{w \in L^\infty(\Omega)} f^*(U, w), && (4.8)
\end{aligned}$$

where

$$\begin{aligned}
f^*(U, w) &\stackrel{\text{def}}{=} \frac{1}{2}||U - z||^2 + \alpha \int_\Omega \phi^*(|\nabla U|, w) && (4.9)\\
&= \frac{1}{2}||U - z||^2 + \frac{\alpha}{2} \int_\Omega \left(w\,(|\nabla u|^2 + \beta^2) + \frac{1}{w} \right).
\end{aligned}$$

The f^* is called the *auxiliary functional*, and w is called the *auxiliary variable.* Note that $f^*(U, w)$ is quadratic in U, and due to (3.8), the inf over w is attained for

$$w^* = \frac{1}{\sqrt{|\nabla U|^2 + \beta^2}}. \tag{4.10}$$

Now alternate minimizations of f^* over U and w: Given an initial guess U^0, for $\nu = 0, 1, \ldots$, set

$$\begin{aligned}
w^{\nu+1} &= \arg\min_w f^*(U^\nu, w) && (4.11)\\
&= \arg\min_w \phi^*(|\nabla U^\nu|, w)\\
&= \frac{1}{\sqrt{|\nabla U^\nu|^2 + \beta^2}}\\
U^{\nu+1} &= \arg\min_U f^*(U, w^{\nu+1}) && (4.12)\\
&= \arg\min_U \frac{1}{2}||U - z||^2 + \frac{\alpha}{2} \int_\Omega \frac{1}{\sqrt{|\nabla U^\nu|^2 + \beta^2}} |\nabla U|^2.
\end{aligned}$$

But this last expression is the variational form for the linear operator equation

$$(I + \alpha L(U^\nu))\, U = z, \tag{4.13}$$

cf., (3.1) and (3.5). Consequently, (4.11)-(4.12) is equivalent to the fixed point iteration (3.10), and the auxiliary variable $w^{\nu+1}$ is the diffusion coefficient for the operator $L(U^\nu)$.

Proceeding as in [4],

$$\begin{aligned}
f(U^{\nu-1}) - f(U^\nu) &= f^*(U^{\nu-1}, w^\nu) - f^*(U^\nu, w^{\nu+1})\\
&= [f^*(U^\nu, w^\nu) - f^*(U^\nu, w^{\nu+1})] + [f^*(U^{\nu-1}, w^\nu) - f^*(U^\nu, w^\nu)].
\end{aligned}$$

By (4.11) and (4.12), both terms in square brackets are nonnegative. Thus $f(U^\nu)$ decreases monotonically, and since it is bounded below by zero, it converges. From this and (4.9),

$$
\begin{aligned}
(4.14)\quad f^*(U^\nu, w^\nu) - f^*(U^\nu, w^{\nu+1}) &= \alpha \int_\Omega \left(\phi^*(|\nabla U^\nu|, w^\nu) - \phi^*(|\nabla U^\nu|, w^{\nu+1})\right) \\
&\downarrow 0 \quad \text{as} \quad \nu \to \infty.
\end{aligned}
$$

Setting $t = |\nabla U^\nu|$ and $s = w^{\nu+1}(x)$ in (4.3), and taking a Taylor expansion in the second variable,

$$
\begin{aligned}
\phi^*(|\nabla U^\nu|, w^\nu) - \phi^*(|\nabla U^\nu|, w^{\nu+1}) &= \frac{\partial \phi^*}{\partial w}(|\nabla U^\nu|, w^{\nu+1})(w^{\nu+1} - w^\nu) \\
&\quad + \frac{1}{2}\frac{\partial^2 \phi^*}{\partial w^2}(|\nabla U^\nu|, \overline{w})(w^{\nu+1} - w^\nu)^2,
\end{aligned}
$$

for some $\overline{w}$ between w^ν and $w^{\nu+1}$. But by (4.11), the first partial derivative is zero. Moreover, by (4.4) and (4.6) the second partial derivative is bounded below by β^3. Thus from (4.14),

$$
(4.15) \qquad \int_\Omega (w^{\nu+1} - w^\nu)^2 \to 0.
$$

Now the gradient of f at $U^{\nu+1}$, cf., (3.3), can be expressed as

$$
(4.16) \qquad g(U^{\nu+1}) = [(I + \alpha L(U^\nu))U^{\nu+1} - z] + \alpha(L(U^{\nu+1}) - L(U^\nu))U^{\nu+1}.
$$

The first term in (4.16) is identically zero by (3.10). From the discussion following (4.12), for any $\varphi \in H^1(\Omega)$,

$$
\begin{aligned}
(4.17)\quad |\langle (L(U^{\nu+1}) - L(U^\nu))U^{\nu+1}, \varphi\rangle| &= |\int_\Omega (w^\nu - w^{\nu-1})\nabla U^{\nu+1} \cdot \nabla \varphi| \\
&\le ||w^\nu - w^{\nu-1}||\, ||U^{\nu+1}||_{H^1(\Omega)}\, ||\varphi||_{H^1(\Omega)}.
\end{aligned}
$$

From Lemma 3.2 and (4.15), the right hand side of (4.17) goes to zero as $\nu \to \infty$. Consequently, the gradient $g(U^\nu)$ converges to zero. Since the functional f is strictly convex, cf., (4.7), we obtain convergence to the minimizer of f. In summary,

THEOREM 4.1. *Consider the discretization (3.7), and let assumptions (**A.1**)-(**A.2**) hold. Then the fixed point iteration (3.10) converges for any initial guess U^0.*

Remark 1. An application closely related to image denoising is the *image deblurring problem.* Here one wishes to recover u from noisy data

$$
z = Ku + \epsilon
$$

where now

$$
(Ku)(x) = \int_\Omega k(x, y)\, u(y)\, dy, \quad x \in \Omega,
$$

and the kernel function k is known. When $k(x,y) = k(x-y)$, the blurring operator K is said to be of convolution form, and the kernel k is called the *point spread function.* The above theorem holds when the fit-to-data term $||U - z||$ in (3.9) is replaced by $||KU - z||$, provided the functional f remains strictly convex and one can establish Lemma 3.2. A sufficient condition to guarantee this is that K is injective.

Remark 2. The uniform bound (3.15)-(3.16) plays a crucial role in establishing convergence, cf., equation (4.17). For typical finite element discretizations with mesh spacing h, $\rho(M_1)$ is proportional to $1/h^2$. Consequently, with increasing discretization, one might expect slower convergence. In the next section, this is verified.

One might also expect a loss of convergence of the fixed point iteration in the continuous setting, provided the minimizer of the functional f is not sufficiently smooth. This difficulty can be overcome by a further modification of the TV penalty functional—for example, by replacing it with

$$\int (\sqrt{|\nabla u|^2 + \beta^2} + \gamma |\nabla u|^2),$$

where γ is another small positive parameter. Chambolle and Lions [3] take a similar approach in their analysis. Modifications such as these ensure both the H^1-boundedness and uniform ellipticity (i.e., H^1-coercivity) of the operators $L(u)$. One can then show in the continuous setting that 2-step iterations like (4.11)-(4.12) are always globally convergent. Unfortunately, the resulting minimizer must lie in $H^1(\Omega)$. This precludes certain desirable phenomena, e.g., discontinuities along edges in 2-D.

5. Local convergence results. In this section we examine (local) rates of convergence, and the dependence of these rates on various parameters. Of particular interest is the "regularization" parameter α in (1.6), which controls the tradeoff between goodness of fit to the data and variability of the solution. Typically this parameter is relatively small—particularly when there is a small amount of error in the data. Also of interest is the "smoothness" parameter β in (1.5). Increasing β has the effect of "rounding off corners" of discontinuous solutions. This parameter is also typically small. The discretization level is another relevant parameter.

For simplicity we assume here that the domain Ω has unit volume. For brevity, we introduce the notation $|\nabla u|_\beta = \sqrt{|\nabla u|^2 + \beta^2}$. Define as before $F(u) = v$ where v is a weak solution to

$$(I + \alpha L(u))v = z \quad \text{in } \Omega, \tag{5.1}$$

$$\frac{\partial v}{\partial \eta} = 0 \quad \text{on } \partial\Omega. \tag{5.2}$$

By standard elliptic theory, F is well-defined as a map from $W^{1,\infty}(\Omega)$ into $H^1(\Omega)$, provided only that $z \in H^1(\Omega)'$. We will, however, assume that $z \in L^2(\Omega)$ in what follows.

Let $\{\phi_j\}_{j=1}^N$ satisfy (**A.1**)–(**A.2**) and assume that the constant function 1 lies in the span of the $\{\phi_j\}$. Consider the "finite element" approximation

$$V = \sum_{j=1}^{N} V_j \phi_j \tag{5.3}$$

to problem (5.1–5.2), which satisfies

$$\int_\Omega V\ \phi_j + \alpha \int_\Omega \frac{\nabla V \cdot \nabla \phi_j}{|\nabla U|_\beta} = \int_\Omega z\ \phi_j \quad \text{for } j = 1, \ldots, N, \tag{5.4}$$

where $U = \sum_{j=1}^{N} U_j \phi_j$. This leads to the natural discretization $F_N : \mathbb{R}^N \to \mathbb{R}^N$ of the iteration map F. Namely, we define $F_N(U) = V$ where V satisfies (5.4). We note that throughout this section, $\{\phi_j\}$ and N are taken to be fixed. As in the previous section, the iterates are bounded. In addition, because of the assumption that $1 \in span\{\phi_j\}$, they preserve the mean of the data (see [15]). More precisely,

Proposition 5.1. *For any initial guess, the iterates $U^{\nu+1} = F_N(U^\nu)$, $\nu = 0, 1, 2, \ldots$, satisfy*

$$||U^{\nu+1}|| \leq ||z||, \tag{5.5}$$

$$\int_\Omega U^{\nu+1} = \int_\Omega z. \tag{5.6}$$

Proof. In (5.4), take $U = U^\nu$ and $V = U^{\nu+1}$. Then replace ϕ_j by $U^{\nu+1}$ and 1, respectively, to obtain (5.5) and (5.6). $\square$

We now establish that the rate of convergence is linear for small values of the parameter α.

Theorem 5.2. *The map F_N is a contraction, provided that the parameter α is sufficiently small. In this case the iterates U^ν defined by $U^{\nu+1} = F_N(U^\nu)$ converge at linear rate in $L^2(\Omega)$ to a unique fixed point, with contraction constant proportional to α.*

Proof. Let K_N be an imbedding constant such that

$$\|\nabla V\|_{L^2} \leq K_N \|V\|_{L^2},$$

for all V of the form (5.3). Let C_N be an imbedding constant such that

$$\|V\|_{L^\infty} \leq C_N \|V\|_{L^2}$$

for all V of the form (5.3). Notice that K_N and C_N both depend on the basis set $\{\phi_j\}$, are finite, and generally grow with N.

To show that F_N is a contraction, consider the difference between successive iterates $V^0 - V^1$, where $V^0 = F(U^0)$ and $V^1 = F(U^1)$. It follows from (5.4) that

$$\begin{aligned} \int |V^0 - V^1|^2 \ + \ \alpha \int \frac{|\nabla V^0 - \nabla V^1|^2}{|\nabla U^1|_\beta} \\ = \ \alpha \int \left(\frac{1}{|\nabla U^0|_\beta} - \frac{1}{|\nabla U^1|_\beta} \right) \nabla V^0 \cdot (\nabla V^1 - \nabla V^0). \end{aligned} \tag{5.7}$$

Hence,

$$\begin{aligned}
\|V^0 - V^1\|_{L^2}^2 &\leq \alpha \left\| \frac{1}{|\nabla U^0|_\beta} - \frac{1}{|\nabla U^1|_\beta} \right\|_{L^\infty} \|\nabla V^0\|_{L^2} \|\nabla V^0 - \nabla V^1\|_{L^2} \\
&\leq \frac{\alpha}{\beta^2} \|\nabla V^0\|_{L^2} \|\nabla V^0 - \nabla V^1\|_{L^2} \|\nabla U^0 - \nabla U^1\|_{L^\infty} \\
&\leq \alpha \frac{C_N K_N^3}{\beta^2} \|V^0\|_{L^2} \|V^0 - V^1\|_{L^2} \|U^0 - U^1\|_{L^2}.
\end{aligned}$$

Since $\|V^0\|_{L^2} \leq \|z\|_{L^2}$, it follows that

$$\|V^0 - V^1\|_{L^2} \leq \alpha \frac{C_N K_N^3}{\beta^2} \|z\|_{L^2} \|U^0 - U^1\|_{L^2}. \tag{5.8}$$

□

Remark 3. It is worth noticing the dependence of the contraction constants on the "smoothing" parameter β. In the estimate (5.8) the term $1/\beta^2$ appears in the contraction constant. Thus the smoother the problem (large β), the faster the convergence. Based on these estimates one might expect a deterioration in the convergence rate as $\beta \to 0$. Numerical experiments such as those in [17] have shown slower convergence of the iteration for small β. Similarly, the appearance of the imbedding constants K_N and C_N indicate that convergence will slow as the discretization level increases. This phenomenon has also been observed in practice.

The idea of the proof of Theorem 5.2 is that for small α, the denoising equation is close to the identity on L^2, so that F_N is a contraction. In a complimentary way, for large α, energy estimates imply that F_N is a contraction in H^1 as we show next.

THEOREM 5.3. *Given a fixed discretization $\{\phi_j\}_{j=1}^N$, the map F_N is a contraction, provided that the parameter α is sufficiently large. In this case the iterates U^ν defined by $U^{\nu+1} = F_N(U^\nu)$ converge at linear rate in $H^1(\Omega)$ to a unique fixed point, with contraction constant proportional to $\alpha^{-1/2}$.*

Proof. Let K_N and C_N be imbedding constants as defined in the proof of Theorem 5.2. From (5.4) we have the energy estimate

$$\int_\Omega |V|^2 + \alpha M \int_\Omega |\nabla V|^2 \leq \int_\Omega |V|^2 + \alpha \int_\Omega \frac{|\nabla V|^2}{|\nabla U|_\beta} \leq \|z\|_{L^2} \|V\|_{L^2}, \tag{5.9}$$

where $M = \inf(1/|\nabla U|_\beta) = 1/\sup |\nabla U|_\beta$. Using (5.5), it then follows that

$$\|\nabla V\|_{L^2} \leq \frac{\|z\|_{L^2}}{\sqrt{\alpha}} (\beta + C_N K_N \|z\|_{L^2})^{1/2} \equiv \frac{R_N \|z\|_{L^2}}{\sqrt{\alpha}} \tag{5.10}$$

Thus the quantity $\|\nabla U^\nu\|_{L^2}$ remains bounded in a ball of radius $R_N \|z\| \alpha^{-1/2}$ for all iterates of the form $U^{\nu+1} = F(U^\nu)$.

From (5.7) and (5.10),

$$
\begin{aligned}
\|\nabla V^0 - \nabla V^1\|_{L^2} &\leq \sup|\nabla U^1|_\beta \left\| \frac{1}{|\nabla U^0|_\beta} - \frac{1}{|\nabla U^1|_\beta} \right\|_{L^\infty} \|\nabla V^0\|_{L^2} \\
&\leq \frac{1}{\beta^2} \sup|\nabla U^1|_\beta \, \|\nabla V^0\|_{L^2} \|\nabla U^0 - \nabla U^1\|_{L^\infty} \\
&\leq \frac{R_N C_N}{\sqrt{\alpha}\beta^2} \|z\|_{L^2} (\beta + C_N K_N \|z\|_{L^2}) \|\nabla U^0 - \nabla U^1\|_{L^2}. \qquad (5.11)
\end{aligned}
$$

So, provided α is large, we see that F_N is a contraction. The convergence of the fixed point iteration in the seminorm $\|\nabla U\|_{L^2}$ follows immediately by the contraction mapping principle. By (5.6) and the Poincare inequality, convergence holds in the H^1 norm. □

Since $\|z\|_{L^2}$ appears in the contraction constant in the estimates in both theorems 5.3 and 5.2, these results taken together imply the following.

COROLLARY 5.4. *If the data z is sufficiently small in the L^2 norm with respect to β then the fixed point map F_N is contraction for every value of α.*

Note, however, that since the contraction constant depends on the ratio $\|z\|/\beta^2$, and since β is usually chosen according to the relative scale of z, one gains neither a guaranteed contraction nor a better convergence rate by rescaling the problem.

Finally, we remark that the difficulty in extending the methods used in this section to the infinite dimensional (non-discrete) setting lies primarily in controlling the L^∞ norm of $|\nabla u|_\beta$, which appears as a coefficient in the operator $L(u)$. Control of $|\nabla u|_\beta$ is necessary to maintain the uniform ellipticity of L. This difficulty can be overcome *on one-dimensional domains* by using the improved regularity of solutions. Contraction mapping estimates can then be obtained similar to those above which, for certain ranges of parameter values, guarantee that the fixed-point iteration converges independent of the discretization level. Of course, control of $|\nabla u|_\beta$ implies smoothness of the solution.

REFERENCES

[1] R. Acar and C. R. Vogel, *Analysis of total variation penalty methods*, Inverse Problems, vol. 10 (1994), pp. 1217-1229.

[2] T. F. Chan, H. M. Zhou, and R. H. Chan, *A Continuation Method for Total Variation Denoising Problems*, UCLA CAM Report 95-18.

[3] A. Chambolle and P. L. Lions, *Image recovery via total variation minimization and related problems*, Research Report No. 9509, CEREMADE, Universite de Paris-Dauphine, 1995.

[4] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, *Deterministic edge-preserving regularization in computed imaging*, Research Report no. 94-01, Univ. of Nice–Sophia Antipolis, 1994.

[5] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems,* North-Holland, Amsterdam (1978).

[6] J. E. Dennis and R. B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice-Hall, Englewood Cliffs NJ (1983).

[7] D. Dobson and F. Santosa, *Recovery of blocky images from noisy and blurred data*, SIAM J. Appl. Math., to appear.

[8] D. Dobson and F. Santosa, *An image enhancement technique for electrical impedance tomography*, Inverse Problems, vol. 10 (1994), pp. 317-334.

[9] D. Geman and G. Reynolds, *Constrained image restoration and the recovery of discontinuities*, PAMI, vol. 14 (1992), pp. 367-383.

[10] D. Geman and C. Yang, *Nonlinear image recovery with half-quadratic regularization*, IEEE Trans. Image Proc., vol. 4 (1995), pp. 932-945.

[11] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Second Edition, Springer-Verlag, Berlin (1983).

[12] S. Gutman, *Identification of discontinuous parameters in flow equations*, SIAM J. Control Optim., vol. 28 (1990), pp. 1049–1060.

[13] K. Ito and K. Kunisch, *An active set strategy for image restoration based on the augmented Lagrangian formulation*, preprint, Center for Research in Scientific Computation, North Carolina State University.

[14] Y. Li and F. Santosa, *An affine scaling algorithm for minimizing total variation in image enhancement*, tech. report CTC94TR201, Cornell Theory Center, Cornell University.

[15] L. I. Rudin, S. Osher, E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D, vol. 60 (1992), pp. 259-268.

[16] L. I. Rudin, S. Osher, C. Fu, *Total Variation Based Restoration of Noisy, Blurred Images*, SIAM J. Numer. Analysis (submitted).

[17] C. R. Vogel and M. E. Oman, *Iterative Methods for Total Variation Denoising*, SIAM J. Sci. Comput., to appear.

[18] C. R. Vogel and M. E. Oman, *Fast Numerical Methods for Total Variation Minimization in Image Reconstruction*, in SPIE Proceedings Vol. 2563, Advanced Signal Processing Algorithms, July, 1995.